

<110> Betenbaugh et al.

<120> Engineering Intracellular Sialylation Pathways

<130> PF509P2

<140> Unassigned

<141> Herewith

<150> 60/227,579

<151> 2000-08-25

<150> 09/516,793

<151> 2000-03-01

<150> 60/169,624

<151> 1999-12-08

<150> 60/122,582

<151> 1999-03-02

<160> 8

<170> PatentIn Ver. 2.1

<210> 1

<211> 1429

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(693)

<400> 1

atg gcc ttc cca aag aag aaa ctt cag ggt ctt gtg gct gca acc atc 48
Met Ala Phe Pro Lys Lys Lys Leu Gln Gly Leu Val Ala Ala Thr Ile
1 5 10 15

acg cca atg act gag aat gga gaa atc aac ttt tca gta att ggt cag 96
Thr Pro Met Thr Glu Asn Gly Glu Ile Asn Phe Ser Val Ile Gly Gln
20 25 30

tat gtg gat tat ctt gtg aaa gaa cag gga gtg aag aac att ttt gtg 144
Tyr Val Asp Tyr Leu Val Lys Glu Gln Gly Val Lys Asn Ile Phe Val
35 40 45

aat ggc aca aca gga gaa ggc ctg tcc ctg agc gtc tca gag cgt cgc 192
Asn Gly Thr Thr Gly Glu Gly Leu Ser Leu Ser Val Ser Glu Arg Arg
50 55 60

cag gtt gca gag gag tgg gtg aca aaa ggg aag gac aag ctg gat cag 240
Gln Val Ala Glu Glu Trp Val Thr Lys Gly Lys Asp Lys Leu Asp Gln
65 70 75 80

gtg ata att cac gta gga gca ctg agc ttg aag gag tca cag gaa ctg 288
Val Ile Ile His Val Gly Ala Leu Ser Leu Lys Glu Ser Gln Glu Leu

	85	90	95	
gcc caa cat gca gca gaa ata gga gct gat ggc atc gct gtc att gca	336			
Ala Gln His Ala Ala Glu Ile Gly Ala Asp Gly Ile Ala Val Ile Ala				
100 105 110				
cgc ttc ttc ctc aag cca tgg acc aaa gat atc ctg att aat ttc cta	384			
Pro Phe Phe Leu Lys Pro Trp Thr Lys Asp Ile Leu Ile Asn Phe Leu				
115 120 125				
aag gaa gtg gct gct gcc gcc cct gcc ctg cca ttt tat tac tat cac	432			
Lys Glu Val Ala Ala Ala Ala Pro Ala Leu Pro Phe Tyr Tyr Tyr His				
130 135 140				
att cct gcc ttg aca ggg gta aag att cgt gct gag gag ttg ttg gat	480			
Ile Pro Ala Leu Thr Gly Val Lys Ile Arg Ala Glu Glu Leu Leu Asp				
145 150 155 160				
ggg att ctg gat aag atc ccc acc ttc caa ggg ctg aaa ttc agt gat	528			
Gly Ile Leu Asp Lys Ile Pro Thr Phe Gln Gly Leu Lys Phe Ser Asp				
165 170 175				
aca gat ctc tta gac ttc ggg caa tgt gtt gat cag aat cgc cag caa	576			
Thr Asp Leu Leu Asp Phe Gly Gln Cys Val Asp Gln Asn Arg Gln Gln				
180 185 190				
cag ttt gct ttc ctt ttt ggg gtg gat gag caa ctg ttg agt gct ctg	624			
Gln Phe Ala Phe Leu Phe Gly Val Asp Glu Gln Leu Leu Ser Ala Leu				
195 200 205				
gtg atg gga gca act gga gca gtg ggc agt ttt gta tcc aga gat tta	672			
Val Met Gly Ala Thr Gly Ala Val Gly Ser Phe Val Ser Arg Asp Leu				
210 215 220				
tca act ttg ttg tca aac tag gttttggagt gtcacagacc aaagccatca	723			
Ser Thr Leu Leu Ser Asn				
225 230				
tgactctggt ctctgggatt ccaatgggcc caccgccgt tccactgcag aaagccatca	783			
gggagtttac tgatagtgtc gaagctaaac tgaagagcct ggatttcctt tctttcactg	843			
atttaaaggga tggaaacttg gaagctggta gctagtgcct ctctatcaaa tcagggtttg	903			
caccttgaga cataatctac cttaaatagt gcattttttt ctacgggaat tttagatgaa	963			
cttgaataaa ctctcctagc aaatgaaatc tcacaataag cattgaggtg cctttttgtg	1023			
gccttaaaaa gtcttatttt gtgaaggggc aaaaactcta ggagtcacaa ctctcagtcg	1083			
ttcattttcag agattttttt gtggagaaat ttctgtttat atggatgaaa tggaaatcaag	1143			
aggaaaattg taattgatta attccatctg tcttttaggag ctctcattat ctccgtctct	1203			
ggttctcaat cctatttttaa agttgtctaa ttttaaacca ctataatatg tcttcatttt	1263			
aataaatatt catttggaa ctaggaaaac tctgagctac tgcatttagg caggcacttt	1323			

aataccaaac tgtaacatgt ctcaactgta tacaactcaa aatacaccag ctcatttggc 1383
 tgctcagtcct aactctagaa tggatgcttt tgaattcatt tcgatg 1429

<210> 2
 <211> 230
 <212> PRT
 <213> Homo sapiens

<400> 2
 Met Ala Phe Pro Lys Lys Lys Leu Gln Gly Leu Val Ala Ala Thr Ile
 1 5 10 15
 Thr Pro Met Thr Glu Asn Gly Glu Ile Asn Phe Ser Val Ile Gly Gln
 20 25 30
 Tyr Val Asp Tyr Leu Val Lys Glu Gln Gly Val Lys Asn Ile Phe Val
 35 40 45
 Asn Gly Thr Thr Gly Glu Gly Leu Ser Leu Ser Val Ser Glu Arg Arg
 50 55 60
 Gln Val Ala Glu Glu Trp Val Thr Lys Gly Lys Asp Lys Leu Asp Gln
 65 70 75 80
 Val Ile Ile His Val Gly Ala Leu Ser Leu Lys Glu Ser Gln Glu Leu
 85 90 95
 Ala Gln His Ala Ala Glu Ile Gly Ala Asp Gly Ile Ala Val Ile Ala
 100 105 110
 Pro Phe Phe Leu Lys Pro Trp Thr Lys Asp Ile Leu Ile Asn Phe Leu
 115 120 125
 Lys Glu Val Ala Ala Ala Ala Pro Ala Leu Pro Phe Tyr Tyr Tyr His
 130 135 140
 Ile Pro Ala Leu Thr Gly Val Lys Ile Arg Ala Glu Glu Leu Leu Asp
 145 150 155 160
 Gly Ile Leu Asp Lys Ile Pro Thr Phe Gln Gly Leu Lys Phe Ser Asp
 165 170 175
 Thr Asp Leu Leu Asp Phe Gly Gln Cys Val Asp Gln Asn Arg Gln Gln
 180 185 190
 Gln Phe Ala Phe Leu Phe Gly Val Asp Glu Gln Leu Leu Ser Ala Leu
 195 200 205
 Val Met Gly Ala Thr Gly Ala Val Gly Ser Phe Val Ser Arg Asp Leu
 210 215 220
 Ser Thr Leu Leu Ser Asn
 225 230

<210> 3
 <211> 1305
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (1)..(1305)

<220>
 <221> Site
 <222> (397)..(399)
 <223> Xaa equals His or Tyr

<220>
 <221> Site
 <222> (407)..(409)
 <223> Xaa equals Gly or Val

<220>
 <221> Site
 <222> (439)..(441)
 <223> Xaa equals Pro or Ser

<220>
 <221> Site
 <222> (505)..(507)
 <223> Xaa equals Gly or Val

<400> 3
 atg gac tcg gtg gag aag ggg gcc gcc acc tcc gtc tcc aac ccg cgg 48
 Met Asp Ser Val Glu Lys Gly Ala Ala Thr Ser Val Ser Asn Pro Arg
 1 5 10 15

ggg cga ccg tcc cgg ggc cgg ccg cgg aag ctg cag cgc aac tct cgc 96
 Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Asn Ser Arg
 20 25 30

ggc ggc cag ggc cga ggt gtg gag aag ccc cgg cac ctg gca gcc cta 144
 Gly Gly Gln Gly Arg Gly Val Glu Lys Pro Pro His Leu Ala Ala Leu
 35 40 45

att ctg gcc cgg gga ggc agc aaa ggc atc ccc ctg aag aac att aag 192
 Ile Leu Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu Lys Asn Ile Lys
 50 55 60

cac ctg gcg ggg gtc ccg ctc att ggc tgg gtc ctg cgt gcg gcc ctg 240
 His Leu Ala Gly Val Pro Leu Ile Gly Trp Val Leu Arg Ala Ala Leu
 65 70 75 80

gat tca ggg gcc ttc cag agt gta tgg gtt tgg aca gac cat gat gaa 288
 Asp Ser Gly Ala Phe Gln Ser Val Trp Val Ser Thr Asp His Asp Glu
 85 90 95

att gag aat gtg gcc aaa caa ttt ggt gca caa gtt cat cga aga agt 336
 Ile Glu Asn Val Ala Lys Gln Phe Gly Ala Gln Val His Arg Arg Ser
 100 105 110

tct gaa gtt tca aaa gac agc tct acc tca cta gat gcc atc ata gaa 384
 Ser Glu Val Ser Lys Asp Ser Ser Thr Ser Leu Asp Ala Ile Ile Glu
 115 120 125

ttt ctt aat tat yat aat gag gkt gac att gta gga aat att caa gct 432
 Phe Leu Asn Tyr Xaa Asn Glu Xaa Asp Ile Val Gly Asn Ile Gln Ala
 130 135 140

act tct yca tgt tta cat cct act gat ctt caa aaa gtt gca gaa atg 480
 Thr Ser Xaa Cys Leu His Pro Thr Asp Leu Gln Lys Val Ala Glu Met
 145 150 155 160

att cga gaa gaa gga tat gat tct gkt ttc tct gtt gtg aga cgc cat 528
 Ile Arg Glu Glu Gly Tyr Asp Ser Xaa Phe Ser Val Val Arg Arg His

165	170	175	
cag ttt cga tgg agt gaa att Gln Phe Arg Trp Ser Glu Ile	cag aaa gga gtt cgt gaa gtg acc gaa Gln Lys Gly Val Arg Glu Val Thr Glu	576	
180	185	190	
cct ctg aat tta aat cca gct Pro Leu Asn Leu Asn Pro Ala	aaa cgg cct cgt cga caa gac tgg gat Lys Arg Pro Arg Arg Gln Asp Trp Asp	624	
195	200	205	
gga gaa tta tat gaa aat ggc Gly Glu Leu Tyr Glu Asn Gly	tca ttt tat ttt gct aaa aga cat ttg Ser Phe Tyr Phe Ala Lys Arg His Leu	672	
210	215	220	
ata gag atg ggt tac ttg cag ggt gga Ile Glu Met Gly Tyr Leu Gln Gly	aaa tgg cat act acg aaa tgc Lys Trp His Thr Thr Lys Cys	720	
225	230	235	
gag ctg gaa cat agt gtg gat Glu Leu Glu His Ser Val Asp	ata gat gtg gat att gat tgg cct att Ile Asp Val Asp Ile Asp Trp Pro Ile	768	
245	250	255	
gca gag caa aga gta tta aga Ala Glu Gln Arg Val Leu Arg	tat ggc tat ttt ggc aaa gag aag ctt Gly Tyr Phe Gly Lys Glu Lys Leu	816	
260	265	270	
aag gaa ata aaa ctt ttg gtt Lys Glu Ile Lys Leu Leu Val	tgc aat att gat gga tgt ctc acc aat Cys Asn Ile Asp Gly Cys Leu Thr Asn	864	
275	280	285	
ggc cac att tat gta tca gga Gly His Ile Tyr Val Ser Gly	gac caa aaa gaa ata ata tct tat gat Asp Gln Lys Glu Ile Ile Ser Tyr Asp	912	
290	295	300	
gta aaa gat gct att ggg Val Lys Asp Ala Ile Gly	ata agt tta tta aag aaa agt ggt att gag Ile Ser Leu Leu Lys Lys Ser Gly Ile Glu	960	
305	310	315	
gtg agg cta atc tca gaa agg Val Arg Leu Ile Ser Glu Arg	gcc tgt tca aag cag acg ctg tct tct Cys Ser Lys Gln Thr Leu Ser Ser	1008	
325	330	335	
tta aaa ctg gat tgc aaa atg Leu Lys Leu Asp Cys Lys Met	gaa gtc agt gta tca gac aag cta gca Glu Val Ser Val Ser Asp Lys Leu Ala	1056	
340	345	350	
gtt gta gat gaa tgg aga Val Val Asp Glu Trp Arg	aaa gaa atg ggc ctg tgc tgg aaa gaa gtg Lys Glu Met Gly Leu Cys Trp Lys Glu Val	1104	
355	360	365	
gca tat ctt gga aat gaa gtg Ala Tyr Leu Gly Asn Glu Val	tct gat gaa gag tgc ttg aag aga gtg Ser Asp Glu Glu Cys Leu Lys Arg Val	1152	
370	375	380	
ggc cta agt ggc gct cct gct Gly Leu Ser Gly Ala Pro	gct gat gcc tgt tcc tac gcc cag aag gct Ala Asp Ala Cys Ser Tyr Ala Gln Lys Ala	1200	
385	390	395	
		400	

gtt gga tac att tgc aaa tgt aat ggt ggc cgt ggt gcc atc cga gaa 1248
 Val Gly Tyr Ile Cys Lys Cys Asn Gly Gly Arg Gly Ala Ile Arg Glu
 405 410 415

ttt gca gag cac att tgc cta cta atg gaa aaa gtt aat aat tca tgc 1296
 Phe Ala Glu His Ile Cys Leu Leu Met Glu Lys Val Asn Asn Ser Cys
 420 425 430

caa aaa tag 1305
 Gln Lys

<210> 4
 <211> 434
 <212> PRT
 <213> Homo sapiens

<220>
 <221> Site
 <222> (133)
 <223> Xaa equals His or Tyr

<220>
 <221> Site
 <222> (136)
 <223> Xaa equals Gly or Val

<220>
 <221> Site
 <222> (147)
 <223> Xaa equals Pro or Ser

<220>
 <221> Site
 <222> (169)
 <223> Xaa equals Gly or Val

<400> 4
 Met Asp Ser Val Glu Lys Gly Ala Ala Thr Ser Val Ser Asn Pro Arg
 1 5 10 15
 Gly Arg Pro Ser Arg Gly Arg Pro Pro Lys Leu Gln Arg Asn Ser Arg
 20 25 30
 Gly Gly Gln Gly Arg Gly Val Glu Lys Pro Pro His Leu Ala Ala Leu
 35 40 45
 Ile Leu Ala Arg Gly Gly Ser Lys Gly Ile Pro Leu Lys Asn Ile Lys
 50 55 60
 His Leu Ala Gly Val Pro Leu Ile Gly Trp Val Leu Arg Ala Ala Leu
 65 70 75 80
 Asp Ser Gly Ala Phe Gln Ser Val Trp Val Ser Thr Asp His Asp Glu
 85 90 95
 Ile Glu Asn Val Ala Lys Gln Phe Gly Ala Gln Val His Arg Arg Ser
 100 105 110
 Ser Glu Val Ser Lys Asp Ser Ser Thr Ser Leu Asp Ala Ile Ile Glu
 115 120 125
 Phe Leu Asn Tyr Xaa Asn Glu Xaa Asp Ile Val Gly Asn Ile Gln Ala
 130 135 140
 Thr Ser Xaa Cys Leu His Pro Thr Asp Leu Gln Lys Val Ala Glu Met
 145 150 155 160

```

Ile Arg Glu Glu Gly Tyr Asp Ser Xaa Phe Ser Val Val Arg Arg His
      165      170      175
Gln Phe Arg Trp Ser Glu Ile Gln Lys Gly Val Arg Glu Val Thr Glu
      180      185      190
Pro Leu Asn Leu Asn Pro Ala Lys Arg Pro Arg Arg Gln Asp Trp Asp
      195      200      205
Gly Glu Leu Tyr Glu Asn Gly Ser Phe Tyr Phe Ala Lys Arg His Leu
      210      215      220
Ile Glu Met Gly Tyr Leu Gln Gly Gly Lys Trp His Thr Thr Lys Cys
      225      230      235      240
Glu Leu Glu His Ser Val Asp Ile Asp Val Asp Ile Asp Trp Pro Ile
      245      250      255
Ala Glu Gln Arg Val Leu Arg Tyr Gly Tyr Phe Gly Lys Glu Lys Leu
      260      265      270
Lys Glu Ile Lys Leu Leu Val Cys Asn Ile Asp Gly Cys Leu Thr Asn
      275      280      285
Gly His Ile Tyr Val Ser Gly Asp Gln Lys Glu Ile Ile Ser Tyr Asp
      290      295      300
Val Lys Asp Ala Ile Gly Ile Ser Leu Leu Lys Lys Ser Gly Ile Glu
      305      310      315      320
Val Arg Leu Ile Ser Glu Arg Ala Cys Ser Lys Gln Thr Leu Ser Ser
      325      330      335
Leu Lys Leu Asp Cys Lys Met Glu Val Ser Val Ser Asp Lys Leu Ala
      340      345      350
Val Val Asp Glu Trp Arg Lys Glu Met Gly Leu Cys Trp Lys Glu Val
      355      360      365
Ala Tyr Leu Gly Asn Glu Val Ser Asp Glu Glu Cys Leu Lys Arg Val
      370      375      380
Gly Leu Ser Gly Ala Pro Ala Asp Ala Cys Ser Tyr Ala Gln Lys Ala
      385      390      395      400
Val Gly Tyr Ile Cys Lys Cys Asn Gly Gly Arg Gly Ala Ile Arg Glu
      405      410      415
Phe Ala Glu His Ile Cys Leu Leu Met Glu Lys Val Asn Asn Ser Cys
      420      425      430
Gln Lys

```

<210> 5

<211> 1080

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (1)..(1080)

<400> 5

```

atg ccg ctg gag ctg gag ctg tgt ccc ggg cgc tgg gtg ggc ggg caa 48
Met Pro Leu Glu Leu Glu Leu Cys Pro Gly Arg Trp Val Gly Gly Gln
      1      5      10      15

```

```

cac ccg tgc ttc atc att gcc gag atc ggc cag aac cac cag ggc gac 96
His Pro Cys Phe Ile Ile Ala Glu Ile Gly Gln Asn His Gln Gly Asp
      20      25      30

```

```

ctg gac gta gcc aag cgc atg atc cgc atg gcc aag gag tgt ggg gct 144
Leu Asp Val Ala Lys Arg Met Ile Arg Met Ala Lys Glu Cys Gly Ala

```

35	40	45	
gat tgt gcc aag ttc cag aag agt gag cta gaa ttc aag ttt aat cgg Asp Cys Ala Lys Phe Gln Lys Ser Glu Leu Glu Phe Lys Phe Asn Arg 50 55 60			192
aaa gcc ttg gag agg cca tac acc tcg aag cat tcc tgg ggg aag acg Lys Ala Leu Glu Arg Pro Tyr Thr Ser Lys His Ser Trp Gly Lys Thr 65 70 75 80			240
tac ggg gag cac aaa cga cat ctg gag ttc agc cat gac cag tac agg Tyr Gly Glu His Lys Arg His Leu Glu Phe Ser His Asp Gln Tyr Arg 85 90 95			288
gag ctg cag agg tac gcc gag gag gtt ggg atc ttc ttc act gcc tct Glu Leu Gln Arg Tyr Ala Glu Glu Val Gly Ile Phe Phe Thr Ala Ser 100 105 110			336
ggc atg gat gag atg gca gtt gaa ttc ctg cat gaa ctg aat gtt cca Gly Met Asp Glu Met Ala Val Glu Phe Leu His Glu Leu Asn Val Pro 115 120 125			384
ttt ttc aaa gtt gga tct gga gac act aat aat ttt cct tat ctg gaa Phe Phe Lys Val Gly Ser Gly Asp Thr Asn Asn Phe Pro Tyr Leu Glu 130 135 140			432
aag aca gcc aaa aaa ggt cgc cca atg gtg atc tcc agt ggg atg cag Lys Thr Ala Lys Lys Gly Arg Pro Met Val Ile Ser Ser Gly Met Gln 145 150 155 160			480
tca atg gac acc atg aag caa gtt tat cag atc gtg aag ccc ctc aac Ser Met Asp Thr Met Lys Gln Val Tyr Gln Ile Val Lys Pro Leu Asn 165 170 175			528
ccc aac ttc tgc ttc ttg cag tgt acc agc gca tac cgg ctc cag cct Pro Asn Phe Cys Phe Leu Gln Cys Thr Ser Ala Tyr Pro Leu Gln Pro 180 185 190			576
gag gac gtc aac ctg cgg gtc atc tcg gaa tat cag aag ctc ttt cct Glu Asp Val Asn Leu Arg Val Ile Ser Glu Tyr Gln Lys Leu Phe Pro 195 200 205			624
gac att ccc ata ggg tat tct ggg cat gaa aca ggc ata gcg ata tct Asp Ile Pro Ile Gly Tyr Ser Gly His Glu Thr Gly Ile Ala Ile Ser 210 215 220			672
gtg gcc gca gtg gct ctg ggg gcc aag gtg ttg gaa cgt cac ata act Val Ala Ala Val Ala Leu Gly Ala Lys Val Leu Glu Arg His Ile Thr 225 230 235 240			720
ttg gac aag acc tgg aag ggg agt gac cac tcg gcc tcg ctg gag cct Leu Asp Lys Thr Trp Lys Gly Ser Asp His Ser Ala Ser Leu Glu Pro 245 250 255			768
gga gaa ctg gcc gag ctg gtg cgg tca gtg cgt ctt gtg gag cgt gcc Gly Glu Leu Ala Glu Leu Val Arg Ser Val Arg Leu Val Glu Arg Ala 260 265 270			816

ctg ggc tcc cca acc aag cag ctg ctg ccc tgt gag atg gcc tgc aat 864
 Leu Gly Ser Pro Thr Lys Gln Leu Leu Pro Cys Glu Met Ala Cys Asn
 275 280 285

gag aag ctg ggc aag tct gtg gtg gcc aaa gtg aaa att ccg gaa ggc 912
 Glu Lys Leu Gly Lys Ser Val Val Ala Lys Val Lys Ile Pro Glu Gly
 290 295 300

acc att cta aca atg gac atg ctc acc gtg aag gtg ggt gag ccc aaa 960
 Thr Ile Leu Thr Met Asp Met Leu Thr Val Lys Val Gly Glu Pro Lys
 305 310 315 320

gcc tat cct cct gaa gac atc ttt aat cta gtg gcc aag aag gtc ctg 1008
 Ala Tyr Pro Pro Glu Asp Ile Phe Asn Leu Val Gly Lys Lys Val Leu
 325 330 335

gtc act gtt gaa gag gat gac acc atc atg gaa gaa ttg gta gat aat 1056
 Val Thr Val Glu Glu Asp Asp Thr Ile Met Glu Glu Leu Val Asp Asn
 340 345 350

cat ggc aaa aaa atc aag tct taa 1080
 His Gly Lys Lys Ile Lys Ser
 355

<210> 6
 <211> 359
 <212> PRT
 <213> Homo sapiens

<400> 6
 Met Pro Leu Glu Leu Glu Leu Cys Pro Gly Arg Trp Val Gly Gly Gln
 1 5 10 15
 His Pro Cys Phe Ile Ile Ala Glu Ile Gly Gln Asn His Gln Gly Asp
 20 25 30
 Leu Asp Val Ala Lys Arg Met Ile Arg Met Ala Lys Glu Cys Gly Ala
 35 40 45
 Asp Cys Ala Lys Phe Gln Lys Ser Glu Leu Glu Phe Lys Phe Asn Arg
 50 55 60
 Lys Ala Leu Glu Arg Pro Tyr Thr Ser Lys His Ser Trp Gly Lys Thr
 65 70 75 80
 Tyr Gly Glu His Lys Arg His Leu Glu Phe Ser His Asp Gln Tyr Arg
 85 90 95
 Glu Leu Gln Arg Tyr Ala Glu Glu Val Gly Ile Phe Phe Thr Ala Ser
 100 105 110
 Gly Met Asp Glu Met Ala Val Glu Phe Leu His Glu Leu Asn Val Pro
 115 120 125
 Phe Phe Lys Val Gly Ser Gly Asp Thr Asn Asn Phe Pro Tyr Leu Glu
 130 135 140
 Lys Thr Ala Lys Lys Gly Arg Pro Met Val Ile Ser Ser Gly Met Gln
 145 150 155 160
 Ser Met Asp Thr Met Lys Gln Val Tyr Gln Ile Val Lys Pro Leu Asn
 165 170 175
 Pro Asn Phe Cys Phe Leu Gln Cys Thr Ser Ala Tyr Pro Leu Gln Pro
 180 185 190
 Glu Asp Val Asn Leu Arg Val Ile Ser Glu Tyr Gln Lys Leu Phe Pro
 195 200 205
 Asp Ile Pro Ile Gly Tyr Ser Gly His Glu Thr Gly Ile Ala Ile Ser

```

      210              215              220
Val Ala Ala Val Ala Leu Gly Ala Lys Val Leu Glu Arg His Ile Thr
225              230              235              240
Leu Asp Lys Thr Trp Lys Gly Ser Asp His Ser Ala Ser Leu Glu Pro
      245              250              255
Gly Glu Leu Ala Glu Leu Val Arg Ser Val Arg Leu Val Glu Arg Ala
      260              265              270
Leu Gly Ser Pro Thr Lys Gln Leu Leu Pro Cys Glu Met Ala Cys Asn
      275              280              285
Glu Lys Leu Gly Lys Ser Val Val Ala Lys Val Lys Ile Pro Glu Gly
      290              295              300
Thr Ile Leu Thr Met Asp Met Leu Thr Val Lys Val Gly Glu Pro Lys
      305              310              315              320
Ala Tyr Pro Pro Glu Asp Ile Phe Asn Leu Val Gly Lys Lys Val Leu
      325              330              335
Val Thr Val Glu Glu Asp Asp Thr Ile Met Glu Glu Leu Val Asp Asn
      340              345              350
His Gly Lys Lys Ile Lys Ser
      355

```

```

<210> 7
<211> 1059
<212> DNA
<213> Escherichia coli

```

```

<220>
<221> CDS
<222> (1)..(1041)

```

```

<400> 7
atg agt aat ata tat atc gtt gct gaa att ggt tgc aac cat aat ggt 48
Met Ser Asn Ile Tyr Ile Val Ala Glu Ile Gly Cys Asn His Asn Gly
      1              5              10              15

agt gtt gat att gca aga gaa atg ata tta aaa gcc aaa gag gcc ggt 96
Ser Val Asp Ile Ala Arg Glu Met Ile Leu Lys Ala Lys Glu Ala Gly
      20              25              30

gtt aat gca gta aaa ttc caa aca ttt aaa gct gat aaa tta att tca 144
Val Asn Ala Val Lys Phe Gln Thr Phe Lys Ala Asp Lys Leu Ile Ser
      35              40              45

gct att gca cct aag gca gag tat caa ata aaa aac aca gga gaa tta 192
Ala Ile Ala Pro Lys Ala Glu Tyr Gln Ile Lys Asn Thr Gly Glu Leu
      50              55              60

gaa tct cag tta gaa atg aca aaa aag ctt gaa atg aag tat gac gat 240
Glu Ser Gln Leu Glu Met Thr Lys Lys Leu Glu Met Lys Tyr Asp Asp
      65              70              75              80

tat ctc cat cta atg gaa tat gca gtc agt tta aat tta gat gtt ttt 288
Tyr Leu His Leu Met Glu Tyr Ala Val Ser Leu Asn Leu Asp Val Phe
      85              90              95

tct acc cct ttt gac gaa gac tct att gat ttt tta gca tct ttg aaa 336
Ser Thr Pro Phe Asp Glu Asp Ser Ile Asp Phe Leu Ala Ser Leu Lys

```

100										105										110										
caa	aaa	ata	tgg	aaa	atc	cct	tca	ggt	gag	tta	ttg	aat	tta	ccg	tat	384														
Gln	Lys	Ile	Trp	Lys	Ile	Pro	Ser	Gly	Glu	Leu	Leu	Asn	Leu	Pro	Tyr															
		115						120					125																	
ctt	gaa	aaa	ata	gcc	aag	ctt	ccg	atc	cct	gat	aag	aaa	ata	atc	ata	432														
Leu	Glu	Lys	Ile	Ala	Lys	Leu	Pro	Ile	Pro	Asp	Lys	Lys	Ile	Ile	Ile															
		130				135					140																			
tca	aca	gga	atg	gct	act	att	gat	gag	ata	aaa	cag	tct	ggt	tct	att	480														
Ser	Thr	Gly	Met	Ala	Thr	Ile	Asp	Glu	Ile	Lys	Gln	Ser	Val	Ser	Ile															
		145			150					155				160																
ttt	ata	aat	aat	aaa	ggt	ccg	ggt	ggt	aat	att	aca	ata	tta	cat	tgc	528														
Phe	Ile	Asn	Asn	Lys	Val	Pro	Val	Gly	Asn	Ile	Thr	Ile	Leu	His	Cys															
				165					170					175																
aat	act	gaa	tat	cca	acg	ccc	ttt	gag	gat	gta	aac	ctt	aat	gct	att	576														
Asn	Thr	Glu	Tyr	Pro	Thr	Pro	Phe	Glu	Asp	Val	Asn	Leu	Asn	Ala	Ile															
			180					185				190																		
aat	gat	ttg	aaa	aaa	cac	ttc	cct	aag	aat	aac	ata	ggc	ttc	tct	gat	624														
Asn	Asp	Leu	Lys	Lys	His	Phe	Pro	Lys	Asn	Asn	Ile	Gly	Phe	Ser	Asp															
		195					200					205																		
cat	tct	agc	ggg	ttt	tat	gca	gct	att	gcg	gcg	gtg	cct	tat	gga	ata	672														
His	Ser	Ser	Gly	Phe	Tyr	Ala	Ala	Ile	Ala	Ala	Val	Pro	Tyr	Gly	Ile															
		210				215					220																			
act	ttt	att	gaa	aaa	cat	ttc	act	tta	gat	aaa	tct	atg	tct	ggc	cca	720														
Thr	Phe	Ile	Glu	Lys	His	Phe	Thr	Leu	Asp	Lys	Ser	Met	Ser	Gly	Pro															
		225			230					235				240																
gat	cat	ttg	gcc	tca	ata	gaa	cct	gat	gaa	ctg	aaa	cat	ctt	tgt	att	768														
Asp	His	Leu	Ala	Ser	Ile	Glu	Pro	Asp	Glu	Leu	Lys	His	Leu	Cys	Ile															
			245						250					255																
ggg	gtc	agg	tgt	ggt	gaa	aaa	tct	tta	ggt	tca	aat	agt	aaa	gtg	ggt	816														
Gly	Val	Arg	Cys	Val	Glu	Lys	Ser	Leu	Gly	Ser	Asn	Ser	Lys	Val	Val															
		260						265					270																	
aca	gct	tca	gaa	agg	aag	aat	aaa	atc	gta	gca	aga	aag	tct	att	ata	864														
Thr	Ala	Ser	Glu	Arg	Lys	Asn	Lys	Ile	Val	Ala	Arg	Lys	Ser	Ile	Ile															
		275					280					285																		
gct	aaa	aca	gag	ata	aaa	aaa	ggt	gag	ggt	ttt	tca	gaa	aaa	aat	ata	912														
Ala	Lys	Thr	Glu	Ile	Lys	Lys	Gly	Glu	Val	Phe	Ser	Glu	Lys	Asn	Ile															
		290				295					300																			
aca	aca	aaa	aga	cct	ggt	aat	ggt	atc	agt	ccg	atg	gag	tgg	tat	aat	960														
Thr	Thr	Lys	Arg	Pro	Gly	Asn	Gly	Ile	Ser	Pro	Met	Glu	Trp	Tyr	Asn															
		305			310					315				320																
tta	ttg	ggt	aaa	att	gca	gag	caa	gac	ttt	att	cca	gat	gaa	tta	ata	1008														
Leu	Leu	Gly	Lys	Ile	Ala	Glu	Gln	Asp	Phe	Ile	Pro	Asp	Glu	Leu	Ile															
				325					330					335																
att	cat	agc	gaa	ttc	aaa	aat	cag	ggg	gaa	taa	tgaga	acaaa	aattattg			1059														

Ile His Ser Glu Phe Lys Asn Gln Gly Glu
340 345

<210> 8

<211> 346

<212> PRT

<213> Escherichia coli

<400> 8

Met Ser Asn Ile Tyr Ile Val Ala Glu Ile Gly Cys Asn His Asn Gly
1 5 10 15
Ser Val Asp Ile Ala Arg Glu Met Ile Leu Lys Ala Lys Glu Ala Gly
20 25 30
Val Asn Ala Val Lys Phe Gln Thr Phe Lys Ala Asp Lys Leu Ile Ser
35 40 45
Ala Ile Ala Pro Lys Ala Glu Tyr Gln Ile Lys Asn Thr Gly Glu Leu
50 55 60
Glu Ser Gln Leu Glu Met Thr Lys Lys Leu Glu Met Lys Tyr Asp Asp
65 70 75 80
Tyr Leu His Leu Met Glu Tyr Ala Val Ser Leu Asn Leu Asp Val Phe
85 90 95
Ser Thr Pro Phe Asp Glu Asp Ser Ile Asp Phe Leu Ala Ser Leu Lys
100 105 110
Gln Lys Ile Trp Lys Ile Pro Ser Gly Glu Leu Leu Asn Leu Pro Tyr
115 120 125
Leu Glu Lys Ile Ala Lys Leu Pro Ile Pro Asp Lys Lys Ile Ile Ile
130 135 140
Ser Thr Gly Met Ala Thr Ile Asp Glu Ile Lys Gln Ser Val Ser Ile
145 150 155 160
Phe Ile Asn Asn Lys Val Pro Val Gly Asn Ile Thr Ile Leu His Cys
165 170 175
Asn Thr Glu Tyr Pro Thr Pro Phe Glu Asp Val Asn Leu Asn Ala Ile
180 185 190
Asn Asp Leu Lys Lys His Phe Pro Lys Asn Asn Ile Gly Phe Ser Asp
195 200 205
His Ser Ser Gly Phe Tyr Ala Ala Ile Ala Ala Val Pro Tyr Gly Ile
210 215 220
Thr Phe Ile Glu Lys His Phe Thr Leu Asp Lys Ser Met Ser Gly Pro
225 230 235 240
Asp His Leu Ala Ser Ile Glu Pro Asp Glu Leu Lys His Leu Cys Ile
245 250 255
Gly Val Arg Cys Val Glu Lys Ser Leu Gly Ser Asn Ser Lys Val Val
260 265 270
Thr Ala Ser Glu Arg Lys Asn Lys Ile Val Ala Arg Lys Ser Ile Ile
275 280 285
Ala Lys Thr Glu Ile Lys Lys Gly Glu Val Phe Ser Glu Lys Asn Ile
290 295 300
Thr Thr Lys Arg Pro Gly Asn Gly Ile Ser Pro Met Glu Trp Tyr Asn
305 310 315 320
Leu Leu Gly Lys Ile Ala Glu Gln Asp Phe Ile Pro Asp Glu Leu Ile
325 330 335
Ile His Ser Glu Phe Lys Asn Gln Gly Glu
340 345